3

ESTIMATING INDIVIDUALIZED RISK OF BREAST CANCER

Mitchell H. Gail[1] and Jacques Benichou[2]

[1]Biostatistics Branch, Division of Cancer Epidemiology and Genetics
National Cancer Institute
6130 Executive Blvd., Room EPN 431
Rockville, MD 20892
[2]University of Rouen Medical School
Biostatistics Unit 1, rue de Germont
76031 Rouen Cedex, France

1. INTRODUCTION

It is estimated[1] that 179,000 women in the United States will be diagnosed with breast cancer in 1998 and that 44,000 will die from it. About one in eight women will develop breast cancer sometime in her life.[2] Many women are therefore concerned about the risk of developing breast cancer, especially women with known risk factors such as a history of breast cancer in close relatives. The purpose of this paper is to review and compare two models for projecting the individualized risk of developing breast cancer over defined age ranges. One model,[3] based on data from the Breast Cancer Detection Demonstration Project (BCDDP), uses information on family history, age at menarche, age at first live birth, number of biopsies, and the presence of atypical hyperplasia. A second model by Claus, Risch, and Thompson,[4] which we refer to as the "Claus model", relies on detailed family history information, including age at breast cancer onset in affected relatives. Published tables[5] based on that model cover individuals with at least one affected relative.

1.1. Absolute Risk versus Relative Risk

It is important to distinguish absolute risk from relative risk. Most studies of risk factors for breast cancer estimate relative risk. Relative risk is the ratio of the age-specific incidence rate among women with specific risk factors to the incidence rate among women without risk factors. For example, using the BCDDP model,[3] one can estimate that a forty-year-old nulliparous woman who began menstruating at age 14,

Cancer Genetics for the Clinician, edited by Shaw.
Kluwer Academic / Plenum Publishers, New York, 1999.

who has had no breast biopsies, and whose mother had breast cancer has a relative risk of 2.76 compared to a forty-year-old woman with no risk factors. Although relative risk estimates are useful for identifying risk factors and for comparing the risk of one woman with that of another, they do not directly measure the chance that a woman will develop breast cancer over a defined age interval.

Absolute risk is the chance that a woman with defined risk factors will develop the disease of interest over a defined age interval. For example, one might want to know the chance that the forty-year-old woman described above would develop breast cancer between ages 40 and 70. From the BCDDP model,[3] one estimates this absolute risk as 0.116 or 11.6%.

Four elements influence the absolute risk of breast cancer. One is the age of the woman. For example, the absolute risk that a thirty-year-old woman who is otherwise like the forty-year-old woman described above would develop breast cancer in the next 30 years is 8.5%, rather than the 11.6% calculated previously. The duration of the age interval is also important. Shorter time periods yield smaller absolute risk. For example, the chance that the thirty-year-old woman above would develop breast cancer in the next ten years is 1.2%. The particular risk factors that a woman has also influence absolute risk. Thus a woman whose risk factors put her at high relative risk will have a higher absolute risk over a given age interval than a woman at lower relative risk. A fourth element that influences absolute risk is the chance of dying of some other disease before the disease of interest develops. These competing risks reduce the absolute risk of breast cancer and can have an appreciable impact in old age. The impact of these four factors is expressed quantitatively in equation (5) of Gail et al.[3] that takes current age, age interval, risk factors, and competing risks into account.

Absolute risk is directly relevant to counseling, because it allows a woman to evaluate the magnitude of the risk over various time periods. An appreciation of absolute risk can lead to a better understanding of the potential benefits of medical options and management strategies. For example, a woman with several risk factors and high relative risk may be reassured to learn that her absolute risk of developing breast cancer over the next ten years is small, and she may elect to follow a program of surveillance. Conversely, she may be very concerned about a large absolute risk over a period of 30 years, and she may decide to undergo prophylactic mastectomy. Such decisions are complicated and depend importantly on the particular concerns of the woman and on the medical options. However, an estimate of absolute risk is a useful ingredient in devising a sound management plan. Information on the range of uncertainty in the estimate of absolute risk is also useful to the woman and health care provider.

1.2. Outline

In Section 2, we review factors associated with increased breast cancer risk. We describe the BCDDP model,[3] easy ways of obtaining absolute risk estimates from this model, and validation studies in Section 3. We review the Claus model[4,5] in Section 4 and compare it to the BCDDP model. In the Discussion (Section 5), we consider how to apply these models in counseling and how to deal with special factors that affect risk, such as demonstration of a mutation in the breast cancer genes, BRCA1 or BRCA2.

Table 1. Selected Risk Factors for Breast Cancer Incidence

Factor	Comparison Group	Approximate Relative Risk	References
Age 60–64	Age 25–29	56	26
Western country	Japan	5	27
Family history of breast cancer			
One first-degree relative	No affected first-degree relative	1.4–3	3, 28
Two or more first-degree relatives		4–6	3
Early age of onset (30 yrs) in an affected relative	Age 50 at onset	2.6	4
BRCA1 or 2 mutation carrier risk (to age 70)	Non-carrier	5–15	6, 7
Age at menarche 11	Age 16	1.3	28
Age at first birth ≥30	Age <20	1.9	28
Age at menopause after 55	Age 45–55	1.5	28
Exposure to 100 rads	No exposure	3	28
Two alcoholic drinks/day	Nondrinker	1.7	28
Hormone replacement therapy ≥10 yrs	None	1.3	28
A breast biopsy	No biopsy or aspiration	1.3–1.7	3, 28
Proliferative disease on biopsy	No biopsy or aspiration	2	28
Atypical hyperplasia on biopsy	No biopsy or aspiration	4	29, 30
≥75% dense tissue on mammogram	No dense tissue	5	31
Contralateral breast cancer	None	5	28

2. RISK FACTORS

Epidemiologists have identified many factors associated with breast cancer risk. It is useful to divide these into factors that may induce cancer and features of the medical history, such as the presence of atypical hyperplasia on a biopsy, that serve as markers of increased risk but may not cause cancer. These two types of risk factors are grouped separately in Table 1.

Age is the most important risk factor; a woman age 70–74 has 56 times the risk of developing breast cancer in the next year as a woman age 25–29 (Table 1). As shown in the solid semi-logarithmic plot in Fig. 1, risk rises with age at a rapid exponential rate for young women and continues to rise exponentially, though at a slower rate, for women over 50 years old. Living in a western country is associated with a relative risk of 5 (Table 1). Women with two or more affected first degree relatives (e.g., mother, sister, or daughter) have high relative risks (4–6), and women known to carry mutations of the breast cancer genes BRCA1 or BRCA2 have cumulative relative risks to age 70 estimated to be between 5 (see Struewing et al.[6]) and 15 (see Whittemore et al.[7]). Relative risks are even higher at younger ages. Aspects of the reproductive history, such as age at menarche, age at first live birth, and age at menopause also affect risk, but the associated relative risks are moderate (Table 1). Recent evidence suggests an association with elevated alcohol intake and with prolonged hormone replacement therapy. Exposure to 100 rads of radiation is associated with a relative risk of 3.

Certain features of the medical history are prognostic, even though they may be markers of the disease process rather than causal agents. Having more than 75% dense tissue on a mammogram carries a relative risk of 5 compared to a woman with no dense

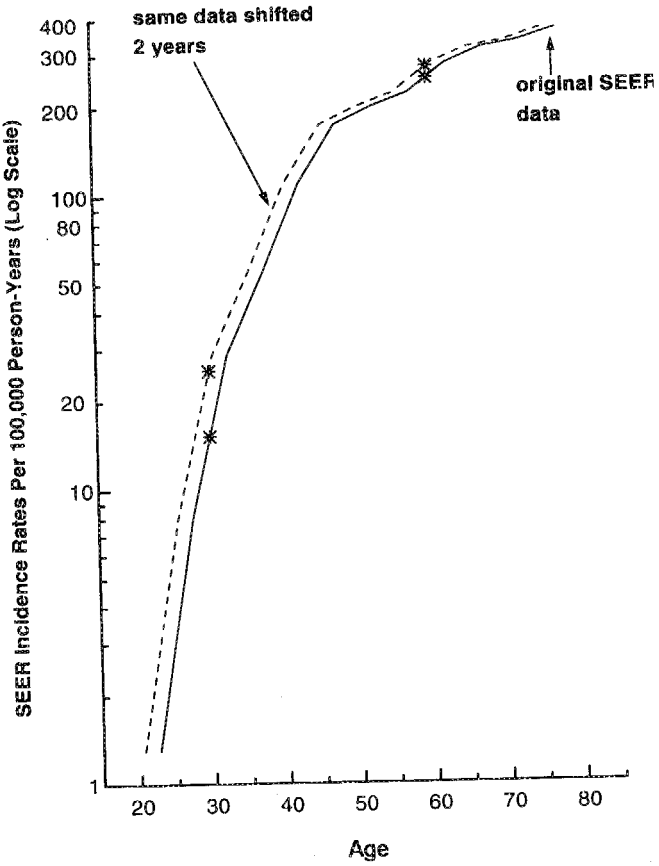

Figure 1. Age-specific breast cancer incidence rates from the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute are shown by the solid curve. Taken from Gail and Benichou.[32] The dashed curve represents the effect of a two-year lead time from screening. Asterisks correspond to relative risks from screening of 1.67 and 1.12 for 30- and 60-year-old women, respectively.

tissue (Table 1). A history of contralateral breast cancer also increases the risk five fold. Demonstration of atypical hyperplasia or proliferative disease on biopsy is associated with elevated risk, and even a history of biopsies with non-proliferative disease is associated with a modest elevation in risk. Such features of the medical history may be useful in projecting risk, even though some of them, such as mammographic density, have not yet been incorporated into models for projecting risk.

It is clear that models for projecting breast cancer risk should at least account for age and family history. Both the BCDDP model[3] and the Claus model[4,5] do this, though the methods used differ.

3. BCDDP MODEL

3.1. Study Population and Modeling Approach

The BCDDP model[3] was based on a study of 243,221 white women who volunteered to be screened annually with mammography and physical exams for five years beginning in 1973–1975. Because the age-specific breast cancer incidence rates are

higher in women in regular screening, Gail et al.[3] noted that "the risk projections from this model are probably most reliable for counseling women who plan to be examined about once a year."

Gail et al. used data from an embedded case-control study to empirically model relative risks from family history and aspects of reproductive and medical history. Although they did not posit a particular genetic model for breast cancer, genetic features are captured to some extent by empirical modeling of risks associated with family history. Gail et al. combined case-control information on relative risks with cohort information on age-specific breast cancer incidence rates to obtain estimates of absolute risk for women with particular risk factors.

3.2. Estimating Risk from the BCDDP Model

The BCDDP model first calculates a multivariate relative risk that takes age at menarche, age at first live birth, number of biopsies, atypical hyperplasia, and number of affected first-degree relatives (mothers or sisters) into account (Table 2). To compute a relative risk from Table 2, one multiplies the four relative risks corresponding to factors A, B, C, and D. For example, the forty-year-old woman considered previously began menstruating at age 14 (relative risk factor 1.00), had no breast biopsies (relative risk factor 1.00), was nulliparous and had an affected mother but no affected sisters (relative risk factor 2.76), and had no atypical hyperplasia (1.00). Thus the combined relative risk is $1.00 \times 1.00 \times 2.76 \times 1.00 = 2.76$. Now consider a 40-year old woman who began menstruating at age 12 (relative risk factor 1.10), and who had one breast biopsy (relative risk factor 1.70). Her first child was born when she was 31 years old, and her mother, but none of her sisters, had a diagnosis of breast cancer (relative risk factor 2.83). There is no information on atypical hyperplasia from the biopsy (relative risk factor 1.00). Her combined relative risk is therefore $1.10 \times 1.70 \times 2.83 \times 1.00 = 5.29$. If atypical hyperplasia had been present the relative risk estimate would be $1.10 \times 1.70 \times 2.83 \times 1.82 = 9.63$.

Gail et al.[3] showed how to convert relative risk estimates into absolute risk estimates using their Table 4 for various ages at counseling and years of follow-up. It is simpler and sufficiently accurate for clinical applications to read absolute risk estimates from the graphs of absolute risk versus relative risk prepared by Benichou, et al.[8] We have reproduced their graphs in Fig. 2 for 30 year risk projections. Panel A is for women with no biopsies, panels B and E for women with one biopsy, and panels C and F for women with more than one biopsy. Separate plots are given for 20-, 30-, 40-, and 50-year-old women. The thirty year risk for the forty-year-old woman with relative risk 2.76 and no breast biopsies can be estimated from Fig. 2A as 12%. From Fig. 2B, one estimates an absolute risk of 17% for the forty year old woman with one biopsy and no information on atypical hyperplasia. From Fig. 2E, the estimate of absolute risk is 28% for the forty year old woman with atypical hyperplasia on her biopsy.

Twenty- and ten-year projections can be obtained from Fig. 2 in Benichou et al.[8] (not shown here). To obtain projections for a 35 year-old woman, one can interpolate between projections for a 30-year-old woman and a 40-year-old woman. Benichou et at.[8] also provide a graph, reproduced as our Fig. 3, of upper and lower confidence level limits plotted against absolute risk. Figure 3 can be used to construct a confidence interval for the risk projection. For example, the woman with an estimated risk of 28% and atypical hyperplasia would have a confidence interval on the risk projection of (17%, 43%). Note from Fig. 3 that the width of the confidence interval increases with

Table 2. Relative Risk Computation for the BCDDP Model. Adapted from Gail et al.[3]

Risk Factor		Relative Risk
A. Age at menarche, years		
≥14		1.00
12–13		1.10
<12		1.21
B. Number of breast biopsies		
Age at counseling < 50 years		
0		1.00
1		1.70
≥2		2.88
Age at counseling ≥ 50 years		
0		1.00
1		1.27
≥2		1.62
	No. of first-degree relatives with breast cancer	
C. Age at first live birth, years		
<20	0	1.00
	1	2.61
	≥2	6.80
20–24	0	1.24
	1	2.68
	≥2	5.78
25–29 or nulliparous	0	1.55
	1	2.76
	2	4.91
≥30	0	1.93
	1	2.83
	≥2	4.17
D. Atypical hyperplasia (AH)		
No biopsies		1.00
At least one biopsy and no AH found on any biopsy		0.93
No AH found and AH status unknown for at least one biopsy		1.00
AH found on at least one biopsy		1.82

To compute overall relative risk, multiply four component risks from categories A, B, C, and D. For example, a 40-year-old nulliparous woman who began menstruating at age 14, who has had no biopsies, and whose mother had breast cancer has an overall relative risk of $1.00 \times 1.00 \times 2.76 \times 1.00 = 2.76$.

increasing estimated absolute risk, reflecting greater uncertainty with higher projected risks.

A simpler and more accurate approach is to use the computer program RISK written by Benichou.[9] This program calculates risk using formulas in Gail et al.,[3] and calculates confidence intervals based on Benichou and Gail[10] and Benichou.[11] Figure 4 depicts an interactive computer dialogue for the forty-year-old woman with atypical hyperplasia and relative risk 9.63. RISK calculates her risk as 29.7%, with 95% confidence interval (19.8%, 41.8%). The previous graphical estimates of 28% with confidence interval (17%, 43%) are close to the exact results from RISK.

3.3. Validation of the BCDDP Model

Several investigators have checked the relative risk calculations of the BCDDP model against other data sources. Because data on atypical hyperplasia are not consis-

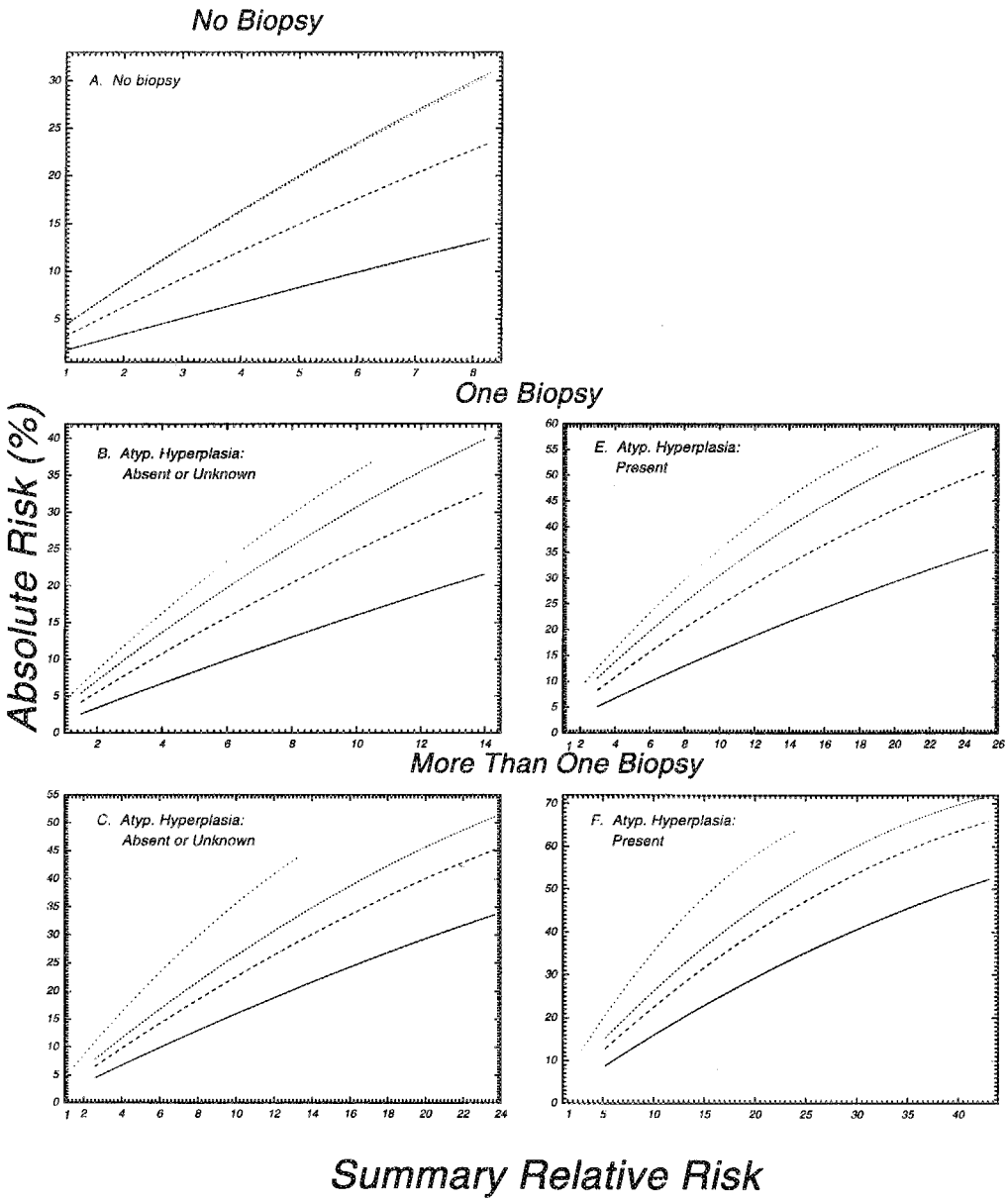

Figure 2. Plots from Benichou et al.[8] of the absolute risk of developing breast cancer in thirty years versus relative risk. Separate curves depict risk for women who are 20 (———), 30 (- - -), 40 (– – –), and 50 (·····) years old at the time of counseling. Figure A corresponds to no biopsies, Figures B and E to one biopsy, and Figures C and F to two or more biopsies.

tently available, only the relative risks associated with age, age at menarche, number of biopsies, number of affected first degree relatives, and age at first live birth have been checked. Gail and Benichou[12] found very similar relative risks when fitting these factors to data from the Cancer and Steroid Hormone (CASH) study. CASH was a population-based case-control study of women between ages 20–54 from the general U.S. population. Cases and controls accrued between December 1, 1980 and Decem-

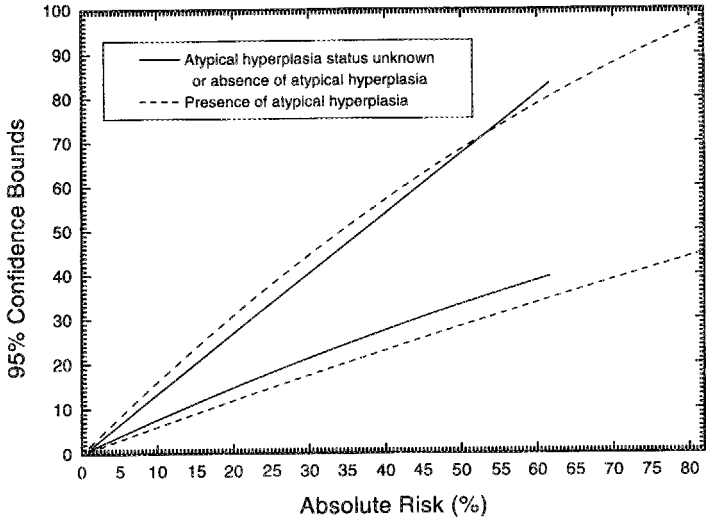

Figure 3. Lower and upper 95% confidence bounds versus estimated absolute risk. Dashed curves apply if atypical hyperplasia has been documented. Otherwise, use solid curves. Taken from Benichou et al.[8]

ber 31, 1982, when there was very little screening mammography. Spiegelman et al.[13] fitted the risk factors in the BCDDP model to data from the Nurses Health Study. Follow-up of this cohort of nurses began in 1976, when the nurses were between the ages of 30 and 55, and follow-up information was used through 1987. Spiegelman et al.[13] found good agreement between the relative risks of the BCDDP model and those obtained by refitting these risk factors to data from the Nurses Health Study, except that relative risks associated with having affected first-degree relatives were smaller.

```
A:\>risk

        Individualized Breast Cancer Risk Projections


  current age (20-80): 40
  upper age limit (20-80): 70

  age at menarche: 12
  age at first live birth (0 if no live birth): 31
  number of previous breast biopsies: 1
  at least one biopsy with hyperplasia
      (y:yes, n:no, u:unknown) ? y
  number of first-degree relatives
      (mother and/or sister(s)) with breast cancer: 1


                  absolute risk = 29.7%

              95% CI = [ 19.8%, 41.8% ]

A:\>
```

Figure 4. Interactive output from RISK[9] for 40 year-old woman with risk profile shown and relative risk 9.63.

Bondy et al.[14] also found consistency between the relative risks from the BCDDP model and those obtained by fitting these risk factors to a cohort of women who participated in the Texas Breast Screening Project and were found to have at least one first-degree relative when screened initially in 1987. Thus, the relative risk portion of the BCDDP model has been shown to apply quite well in three independent groups of women.

The situation is more complicated for absolute risk projection. In their original report,[3] Gail et al. warned that the BCDDP model would overestimate risk in younger women who are not being screened annually. Gail and Benichou[12] noted that the breast cancer rate found in the BCDDP population below age 55 was about 1.62 times that in the general CASH population and that a higher prevalence of risk factors in the BCDDP population could only account for a ratio of 1.15. They concluded that the BCDDP model would overestimate risk in unscreened younger women, such as those in the CASH population, by about $\{1.62/1.15) - 1\} \times 100 = 41\%$.

Why should screening have such an impact on breast cancer rates for younger women but not older women? Mammographic screening allows one to look about 2 years into the future for a young woman. The amount by which one looks into the future is called the "lead time" of the screening test. Thus, a screened 30-year-old woman has the risk of a 32-year-old woman. Figure 1 illustrates the effect of this two-year lead time (see the dotted curve). Because this is a semi-logarithmic plot with a steep slope at younger ages, even a two-year shift increases risk sharply. For example the relative risk for a 30-year-old screened woman, compared to a 30-year-old unscreened woman, is 1.67. For a 60-year-old woman, the relative risk associated with a two-year lead time is only 1.12, because the slope of the age-specific log-incidence curve is much shallower in older women. Screening effects could easily explain much of the discrepancy that arises when the BCDDP model, which was derived from women in regular annual screening, is applied to unscreened or sporadically screened younger women.

Bondy et al.[14] found that the BCDDP model overestimated risk in younger but not older women in the cohort derived from the Texas Breast Screening Project. For women under age 60, the ratio of expected breast cancers under the BCDDP model to observed cancers was $30.58/12 = 2.55$. For women age 60 and over, the ratio was $20.74/25 = 0.83$. The overall ratio of expected to observed cancers was 1.32. When, however, the BCDDP model was applied to the subset of women who adhered to American Cancer Society screening guidelines, the overall ratio of expected to observed cancers was 0.89, with 95% confidence interval 0.62–1.33. Bondy et al. concluded: "Overall, the Gail et al. model accurately predicts risk in women with a family history of breast cancer and who adhere to American Cancer Society guidelines. The model should be used as it was intended, for women who receive annual mammograms."

Speigelman et al.[13] found expected to observed ratios of 1.48 for women under age 50 and 1.16 for women age 50 or more. The discrepancies were greatest in the years 1976–1981, during which mammographic screening was rarely performed, and less during 1982–1987, when there was sporadic screening in the Nurses Health Study cohort.

Another good opportunity to validate the BCDDP model will arise when data from the Breast Cancer Prevention Trial are fully analyzed. This study recently announced that tamoxifen reduced the incidence of breast cancer, compared to placebo, in women above age 59 or in younger women whose risk factors put them at a risk equivalent to that of an average 60-year-old woman.[15] Because women in this

study are screened annually, the control arm will be useful for assessing predictions from the BCDDP model. Preliminary unpublished analyses indicate that the BCDDP model does, in fact, predict absolute risk well in this population.

To summarize, absolute predictions from the BCDDP model may very well be accurate for women in regular screening, but the BCDDP model overestimates risk in younger women who are not screened regularly. Relative risk estimates from the BCDDP model are applicable to screened and unscreened populations.

4. THE AUTOSOMAL DOMINANT MODEL OF CLAUS ET AL.[4]

4.1. The Claus Model

Claus et al.[4] performed segregation analyses on data from relatives of cases and controls in the CASH study. They concluded that an autosomal dominant model for a single breast cancer gene accounted for familial aggregation of disease better than other models. We now know that there are at least two important breast cancer genes, BRCA1 and BRCA2. Nonetheless, the simple autosomal model used by Claus et al.[4] captures much of the useful prognostic information in family history. Claus et al. assumed that carriers of a dominant breast cancer gene would have one age-specific breast cancer incidence rate function (or cumulative risk function) and non-carriers would have another lower age-specific breast cancer incidence rate function. Even though no genotyping was performed, Claus et al. were able to estimate the frequency of the mutant allele as 0.0033 and the cumulative risk functions for carriers and non-carriers. They estimated a cumulative risk to age 70 of 67% for carriers and 5% for non-carriers.

Claus et al. found that the relative risk of cancer comparing carriers with non-carriers is very large in young women (e.g., 43 for ages 30–39) but smaller in older women (e.g., 5 for ages 70–79). If a relative develops breast cancer while young, she is therefore more likely to be a mutation carrier, which increases the chance that the woman being counseled is a carrier. Thus Claus et al. use not only the number of affected relatives but their ages at disease onset.

Because the CASH study recruitment took place in 1980–1982, at which time family history was elicited, the data used to fit the Claus model represent an unscreened general population. It can be anticipated, therefore, that risk projections for younger women from the Claus model will often be lower than projections from the BCDDP model.

Claus et al.[5] present tables for projecting risk according to the age of the woman being counseled and the ages at onset in affected relatives. They give tables for one affected first-degree relative, one affected second-degree relative, two affected first-degree relatives, an affected mother and maternal aunt, an affected mother and paternal aunt, one affected maternal and one paternal second-degree relative, and two affected maternal or paternal second-degree relatives.

Consider again the 40-year-old woman who began menstruating at age 14, who had no breast biopsies, and whose mother developed breast cancer (see Section 3.2). Her 30 year risk was estimated from the BCDDP model as 11.6%. Table 3, which is taken from Table 2 in Claus et al.,[5] describes the cumulative probability of breast cancer for such a woman according to her age and the age at onset in a first-degree relative. Suppose the mother developed breast cancer at age 63. According to Table 3, the

Table 3. Predicted Cumulative Probability of Breast Cancer for a Woman Who Has One First-Degree Relative Affected With Breast Cancer, by Age of Onset of the Affected Relative Taken from Claus et al.[5]

Age of Woman (yr)	First-Degree Relative with Age of Onset (yr)					
	20–29	30–39	40–49	50–59	60–69	70–79
29	0.007	0.005	0.003	0.002	0.002	0.001
39	0.025	0.017	0.012	0.008	0.006	0.005
49	0.062	0.044	0.032	0.023	0.018	0.015
59	0.116	0.086	0.064	0.049	0.040	0.035
69	0.171	0.130	0.101	0.082	0.070	0.062
79	0.211	0.165	0.132	0.110	0.096	0.088

cumulative risk of breast cancer in the woman being counseled is 0.006 through age 39 and 0.070 through age 69. Therefore, the 30 year risk for this 40-year-old woman is (see the formula in Claus et al.[5]) (0.070–0.006)/(1–0.006) = 0.064. If the mother had developed breast cancer at age 28, the risk would be (0.171–0.025)/(1–0.025) = 0.150 or 15%.

No validation studies of the Claus model have been published. No confidence intervals have been presented for projections under this model. This model also does not take competing risks into account, but such risks would have only a small impact for projections to age 60. For projections to age 80, however, competing risks would tend to reduce absolute risk of breast cancer below the values in Table 3.

4.2. Comparison of BCDDP and Claus Models

As the previous example illustrates, it may be difficult to compare projections from the BCDDP and Claus models because they employ different risk factors. In particular, the Claus model does not include reproductive factors or information on biopsies, while the BCDDP model does not use data on age at onset in affected relatives. The BCDDP model incorporated family history information empirically based on logistic regression, and it was found that age at onset in affected relatives did not add useful information to logistic models that contained all the other factors in Table 2. This observation does not imply that age at onset is not useful for predicting risk in other models, such the Claus model.

The tables given by Claus et al.[5] only pertain to counselees with at least one affected relative, whereas the BCDDP model also makes projections for women without affected relatives.

Some comparisons may be drawn for women with given family histories by considering BCDDP projections in the absence of other risk factors and in the presence of all other risk factors except atypical hyperplasia (Table 4). Projections for the Claus model are made assuming that the relatives are affected either at age 50 or at age 20. No projections are available for the Claus model with no affected relatives (Table 4). For 1 affected relative, projections of risk for ages 30 to 60 range from 8% to 22% for the BCDDP model and from 5% to 11% for the Claus model. Although these projections overlap, the BCDDP projections tend to be higher. This probably reflects the lack of screening in the CASH population and the fact that the projections are for a 30-year-old woman. For a 60 year old woman, the corresponding 20 year risks are 3.0%

Table 4. Comparison of Estimated Absolute Risks from Age 30 to Age 60 from the BCDDP and Claus Models

	BCDDP		Claus Model	
Number of affected first-degree relatives	No other factors	All other factors[a]	Affected at age 50	Affected at age 20
0	3%	13%	NA	NA
1	8%	22%	5%	11%
2	20%	35%	13%	28%

[a]Assumes biopsy status for atypical hyperplasia is unknown.

and 25% for the BCDDP model and 6.4% and 10.7% for the Claus model. With two affected first degree relatives (Table 4), there is considerable overlap between projections under the two models, but again results from the BCDDP model tend to be higher.

McGuigan et al.[16] applied both models to 111 women with at least one affected relative in a high risk breast cancer clinic at the University of California at Los Angeles. They estimated the risk from the woman's current age to age 80. Most risks from the BCDDP model were under 20%, the highest risk was 51%, and a typical risk projection of 15% from the BCDDP model corresponded to a risk of about 10% for the Claus model. There were, however, 11 cases in which the Claus model yielded risks above 25% that were more than twice as large as the BCDDP model projections, presumably because these women had relatives with early-onset disease.

5. DISCUSSION

We have reviewed the BCDDP model and Claus model and indicated how to use them. McTiernan et al.[17] compared these methods with earlier life-table projections based on smaller cohorts of relatives of women with breast cancer[18,19] and with a method that classifies a woman into one of four relative risk categories but does not specify absolute risk.[20]

In using the BCDDP model, it is important to consider the nature of the counselee. Projections are likely to be most accurate if she is white, has just been screened and found to be free of disease, and plans to continue in a program of annual follow-up with screening mammography. As we have indicated, the BCDDP model overestimates risk in young women who do not receive annual mammography. The model was fitted to data on white women because there were too few black or Asian women in the BCDDP to obtain reliable estimates.

The data used to fit the Claus model derived from unscreened women. It is possible, therefore, that projections from this model underestimate risk when applied to younger women in regular follow-up, but no validation studies have been reported.

Neither of these models take certain special features into account, and the counselor has an indispensable role in determining the applicability of the model. For example, if the woman is a recent immigrant from rural Japan or China, her relative risk may be reduced by a factor of five (Table 1). If she has had a diagnosis of breast cancer in the contralateral breast, her relative risk is increased five fold. One may wish to modify risk projections to take such factors in Table 1 into account by multiplying the relative risk from the new factor times the relative risk calculated from Table 2 before using the graphs of Benichou et al.[8] to obtain modified BCDDP projections. The

counselor must, at a minimum, interpret projections from the BCDDP model and the Claus model in light of information not included in these models that can have an important modifying influence.

A nice feature of the BCDDP model is that estimates of uncertainty are expressed in the form of confidence intervals. These confidence intervals tend to be wider for larger estimated risks (Fig. 3). Although these confidence intervals measure uncertainty from random variation due to limitations of sample size, they do not reflect systematic errors that could arise, for example, when applying the model to an unscreened population or to a woman with a previous history of breast cancer or previous exposure to radiation.

Exciting progress has been made in identifying the major breast cancer susceptibility genes, BRCA1 and BRCA2. Whittemore et al.[7] estimated the chance that a BRCA1 mutation carrier would develop breast cancer by age 70 at 69%, whereas Ford et al.[21] estimated 87%. Results in Claus et al.,[4] who implicitly studied BRCA1, BRCA2, and possibly other genes, yield cumulative risk to age 70 of 67% for such carriers. Struewing et al.[6] estimated 56% for carriers of three mutations in BRCA1 and BRCA2 among Ashkenazi Jews. To some extent, risks from BRCA1 or BRCA2 mutations are reflected in the family history data in the BCDDP model and Claus model. Nonetheless, if a woman is known to carry a cancer-causing mutation in BRCA1 or BRCA2, that information should be used to project risk. This information could become available through genotyping several members of a highly affected family. If women with breast cancer share the same mutation, that mutation is very likely to confer risk.

Very few women in the general population carry mutations of BRCA1 or BRCA2. Data in Claus et al.[4] suggest a carrier frequency of 0.7% for BRCA1 or BRCA2 mutations, and data for BRCA1 mutations alone suggest carrier frequencies ranging from 0.1 percent[22] to 0.3 percent.[7] Thus, only a very small fraction of women in the general population carry disease-conferring mutations of BRCA1 or BRCA2, although the carrier frequency may approach 2% in Ashkenazi Jews.[6] Even if a woman is found to carry a mutation in BRCA1 or BRCA2, it is not certain that the mutation confers increased risk unless it has been previously identified as a risk-producing allele or unless it is linked with disease in relatives.

Thus, models such as the BCDDP model and Claus model remain useful for the vast majority of women for whom no disease-causing mutation has been identified. Nonetheless the counselor will want to use available information on mutations of BRCA1 and BRCA2 and on other inherited syndromes that confer greatly increased breast cancer risk. The Li-Fraumeni cancer family syndrome, particularly when supported by evidence of a shared mutation in the p53 gene, is indicative of high breast cancer risk.[23] The Cowden multiple hamartoma syndrome is a rare autosomal dominant condition that also confers greatly increased breast cancer risk.[24] These syndromes are exceedingly rare, however.

Ongoing work by several investigators may lead to improved models. One line of research would incorporate additional strong risk factors, such as the presence of dense tissue on mammograms. Barry et al.[25] use family history to estimate the probability that a woman carries a mutation in BRCA1 or BRCA2; this approach may allow one to summarize complex family history data in risk models that also include other risk factors.

An estimate of absolute risk from these models is a useful tool in the counseling process, but it is only one element. The counselor must gather the best available prog-

nostic information, including a careful personal and family history and results of histopathologic examinations, before using such models and must be alert to features not accounted for in the models. Most important, the counselor will need time to put risk estimates, and their uncertainties, into perspective for the counselee and to carefully explain management options in the context of estimated risks. Giving risk projections for several age intervals may be useful. Only by working closely with the counselee to convey information and support and to understand the counselee's preferences and reactions to such information, can the counselor use these models effectively.

REFERENCES

1. Landis SH, Murray T, Bolden S, et al: Cancer Statistics, 1998. CA-A Cancer Journal for Clinicians 48:6–29, 1998
2. Ries LAG, Kosary CL, Hankey BF, et al: SEER Cancer Statistics Review, 1973–1994 Bethesda, MD, National Cancer Institute, NIH Pub. No. 97-2789, page 131, 1997
3. Gail MH, Brinton LA, Byar DP, et al: Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 81:1879–1886, 1989
4. Claus EB, Risch NJ, Thompson WD: Genetic analysis of breast cancer in the Cancer and Steroid Hormone Study. Am J Hum Genet 48:232–242, 1991
5. Claus EB, Risch N, Thompson WD: Autosomal dominant inheritance of early onset breast cancer: Implications for risk prediction. Cancer 73:643–651, 1994
6. Struewing JP, Hartge P, Wacholder S, et al: The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. N Engl J Med 336:1401–1408, 1997
7. Whittemore AS, Gong G, Itnyre J: Prevalence and contribution of BRCA1 mutations in breast cancer and ovarian cancer: Results from three U.S. population-based case-control studies of ovarian cancer. Am J Hum Genet 60:496–504, 1997
8. Benichou J, Gail MH, Mulvihill JJ: Graphs to estimate an individualized risk of breast cancer. J Clin Oncol 14:103–110, 1996
9. Benichou J: A computer program for estimating individualized probabilities of breast cancer. Comput Biomed Res 26:373–382, 1993
10. Benichou J, Gail MH: Methods of inference for estimates of absolute risk derived from population-based case-control studies. Biometrics 51:182–194, 1995
11. Benichou J: A complete analysis of variability for estimates of absolute risk from a population-based case-control study on breast cancer. Biometric J 37:3–24, 1995
12. Gail MH, Benichou J: Assessing the risk of breast cancer in individuals, in DeVita VT, Hellman S, Rosenberg SA (eds): Cancer Prevention. Philadelphia, PA, Lippincott, pp 1–15, 1992
13. Spiegelman D, Colditz GA, Hunter D, et al: Validation of the Gail et al. model for predicting individualized breast cancer risk. J Natl Cancer Inst 86:600–607, 1994
14. Bondy ML, Lustbader ED, Halabi S, et al: Validation of a breast cancer risk assessment model in women with a positive family history. J Natl Cancer Inst 86:600–607, 1994
15. Redmond CK, Costantino JP: Design and current status of the NSABP breast cancer prevention trial, in Senn HJ, Gelber RD, Goldhirsch A, Thurlimann B (eds): Recent Results in Cancer Research 140: Adjuvant Therapy of Breast Cancer. Berlin, Springer-Verlag, pp 309–317, 1996
16. McGuigan KA, Ganz PA, Breant C: Agreement between breast cancer risk estimation methods. J Natl Cancer Inst 88:1315–1317, 1996
17. McTiernan A, Gilligan MA, Redmond C: Asessing individual risk for breast cancer: Risky business. J Clin Epidemiol 50:547–556, 1997
18. Ottman R, Pike MC, King MC, et al: Practical guide for estimating risk for familial breast cancer. Lancet *ii*:556–558, 1983
19. Anderson DE, Badzioch MD: Risk of familial breast cancer. Cancer 56:383–387, 1985
20. Taplin S, Thompson RS, Schnitzer F, et al: Revisions in the risk-based breast cancer screening program at Group Health Cooperative. Cancer 66:812–818, 1990
21. Ford D, Easton DF, Bishop DT, et al: Risk of cancer in BRCA-1 mutation carriers. Lancet 343:692–695, 1994

22. Ford D, Easton DF, Peto J: Estimates of the gene frequency of BRCA1 and its contribution to breast and ovarian cancer incidence. Am J Hum Genet 57:1457–1462, 1995
23. Li F: Familial aggregation, in Schottenfeld D, Fraumeni JF, Jr (eds): Cancer Epidemiology and Prevention (ed 2). New York, Oxford University Press, pp 546–558, 1966
24. Hanssen AMN, Fryns JB: Cowden syndrome. J Med Genet 32:117–119, 1995
25. Berry DA, Parmigianni G, Sanchez J, Schildkraut J, Winer E: Probability of carrying a mutation of breast-ovarian cancer gene BRCA1 based on family history. J Natl Cancer Inst 89:227–228, 1997
26. Miller BA, Ries LAG, Hankey BF, et al: Cancer statistics review: 1973–1989. Natl Cancer Inst. NIH Publ. No. 92-2789, Washington, DC, US Government Printing Office, 1992
27. Kelsey JL: A review of the epidemiology of breast cancer. Epidemiol Rev 1:74–109, 1979
28. Harris JR, Lippman ME, Veronesi U, et al: Breast cancer. N Engl J Med 327:319–328, 1992
29. Dupont WD, Page DL: Risk factors for breast cancer in women with proliferative disease. N Engl J Med 312:146–151, 1985
30. Carter CL, Corle DK, Micozzi MS, et al: A prospective study of breast cancer in 16,692 women with benign breast disease. Am J Epidemiol 128:467–477, 1988
31. Byrne C, Schairer C, Wolfe J, et al: Mammographic features and breast cancer risk: Effects with time, age, and menopause status. J Natl Cancer Inst 87:1622–1629, 1995
32. Gail MH, Benichou J: Validation studies on a model for breast cancer risk. J Natl Cancer Inst 86:573–575, 1994 (editorial)